Optimization-based Road Curve Fitting

Sheng Zhao and Jay A. Farrell

Abstract—Various advanced driver assistance systems (ADAS) are under development that intend to provide improved road safety. These systems require precise road models. In particular, accurate curvature is important for some ADAS applications such as curve over speed and lane departure warning. Existing road models often employ spline functions that are fit by least squares to roadway position data. The curvature calculated for such spline curves may not accurately reflect the curvature of the underlying roadway. This article addresses this problem in an unified framework, using optimization with $\ell_1$-norm regularization. In this approach, known roadway characteristics can be enforced optimally with respect to a cost function which finds the best tradeoff between the match to the available data and the number of changes in curvature. Experimental results with show that the proposed method chooses a sparse set of curvature switching points (i.e., piecewise constant curvature) and achieves a high accuracy fit to the roadway dataset.

I. INTRODUCTION

In advanced driver assistance systems (ADAS) or Vehicle Assist and Automation (VAA) [1], a digital map with a precise model of the road is necessary. ADAS combines vehicle state information [2] with trajectory information from the digital map [3] to provide the driver with important information: over-safe-speed warnings when driving on curved roads [4], road/lane departure warnings and so forth. These applications require the map to have lane-level accuracy (pos. error < 0.5 m) and contain a precise model of the road in the sense that attributes of the road, such as the curvature, are accurate [3], [5]. It is easy to see the importance of roadway position accuracy for ADAS applications, while the accuracy requirement for the other attributes is less obvious. Accurate curvature is necessary to compute the maximum safe speed, for different vehicle classes, on curved roads for driver assistance or autonomous cruise control systems. Moreover, curvature can be used to adjust the beam pattern of the vehicle headlamps to minimize the driver’s lost visual field caused by road curvature [6]. Inaccuracy of the curvature, in particular, incorrect sign of the curvature dramatically affects the performance of such applications. Therefore, the shape of the roadway, in particular, its curvature must be accurately represented; in particular, straight line connections between nodes is not adequate.

Current digital maps, developed to support routing applications, are not accurate enough to support ADAS applications. Accurate road shape and curvature, which were not high priority specifications in the current generation of maps [4], are vital attributes for future generations of maps. This paper focuses on the reconstruction of the road trajectory to achieve high accuracy position and curvature. The overall roadway trajectory fit procedure involves several steps:

1) Data Acquisition: An instrumented car drives along the roadway storing roadway relevant information, e.g., inertial, GNSS, imagery, LIDAR.
2) Smoothing: This is a post-processing step that combines the time-stamped roadway information to compute optimal estimates of positions along, for example, each lane edge or centerline. The result is an ordered, closely spaced, sequence of points that represent a curve of interest for the map. Herein, we represent such a data set as

$$ D = \{(t_k, N_k, E_k, D_k)\}_{k=0}^{N} $$

where $k$ counts over the samples, $t_k$ is the measurement time, and $(N, E, D)$ denote the north, east, and down coordinates in a local tangent plane.
3) Road Curve Fitting: While $D$ represents the curve, it is not convenient for storage. Due to the extremely large cumulative number of lane miles to be stored, more efficient data storage representations are desirable. This step converts the data sequence $D$ into an analytic curve $\gamma(s)$ that is convenient both for storage and for realtime applications.
4) Storage: The parameters of the analytic curve fit $\gamma(s)$ are stored in a GIS database along with representations of other lanes, roads, and roadway features.
5) Real-time Application: Based on a region of interest, GIS database queries extract curves $\gamma(s)$ necessary for realtime applications. The applications may use the analytic representations in various ways: graphic displays, user warnings, search for nearest trajectory point, etc.

The focus of this paper is on Step 3, the problem of road curve fitting.

II. TECHNICAL PROBLEM STATEMENT

The roadway data curve fit is subject to constraints derived from road design rules [7], [8]. One rules is that the road is composed of three basic elements [7]: straight lines, circular curves and transition curves. If we consider straight segments as a special case of circular curves with curvature equal to zero, then the road only consist of circular curves and transition curves. Transition curves are used to connect two circular curves and are often designed to have constant curvature change [9]. Transition curves can either be explicitly modeled as in [7] or approximated as circular curves as in [4]. As argued in [4], the constant curvature approximation
is reasonable because the length of the transition curves is often small. This paper follows the convention, as in [4], to model the road as a curve with segments of piecewise constant curvature. Moreover, while the real road is three-dimensional, the two dimensional flat planar assumption is often used for each local road segment. This paper also assumes that the road is locally two-dimensional.

The problem of road curve fitting can be stated as follows: Given a set of position data \( D = \{N_k, E_k\}_{k=0}^N \) representative of a lane centerline, fit a piecewise \( C^2 \) curve, with piecewise constant curvature to the data.

Extension to three dimensional data and to lane edges is straightforward.

III. LITERATURE REVIEW

The existing methods to curve fit roadways fall into two main categories: spline based curve fitting [3], [8] and parametric curve fitting [4], [7]. In spline based methods, the user selects a spline basis function, then selects the spline parameters to fit the position data. Road attributes such as the tangent and curvature are computed, from the first and second derivatives of the spline by the application software. There are a few difficulties with the spline approach. First, the spline argument is the arclength \( s \) and the arclength in not known part of the data set \( D \). Instead, arclength is estimated from the data. Second, the spline curve fit does not naturally enforce the piecewise constant curvature constraint. Piecewise constant curvature can be attained by selecting splines of the appropriate order. For splines, the points of curvature change are defined by the knot locations; therefore, the knot locations would need to be optimally selected. Therefore, a preprocessing step may be required to estimate arclength and curvature so that the knots can be allocated as a function of arclength at points of curvature discontinuity. Nevertheless, the curvature derived from curve fit splines has high variance that is not related to the actual roadway shape [3], [4]. Third, the spline fitting procedure is not robust to errors in the position data [3], [4]. In contrast to the spline based methods, recent papers [4], [7] report an alternative method based on parametric curve fitting. This method first computes the local pointwise curvature using three consecutive points. Then it segments the road by searching for curvature discontinuities. Finally, each road segment is fit to a parameterized circle function. Neither of the above procedures are optimal. Each has various ad-hoc aspects.

In [10], the authors formulate a related trajectory generation problem as an optimization problem with \( \ell_1 \)-norm regularization. More details of the \( \ell_1 \)-norm regularization approach can be found in [11]. The approach of [10] is to formulate a commanded vehicle trajectory as the output of a linear system with input \( u \). The constraints are that the trajectory must pass near a set of \( M \) points at specified time instants. The optimization is over the magnitude of impulses, defined on a time grid, for the \( p \)-th derivative of \( u \). The \( \ell_1 \)-norm regularization yields a sparse vector of the \( p \)-th derivative of \( u \).

This paper also follows an \( \ell_1 \)-norm regularization approach to proposes an optimization formulation of the roadway curve fitting problem in an unified framework that jointly computes the arclength, finds the curvature switching points, and fits the road segments to \( D \). In contrast to [10], the model is nonlinear, the independent variable is arclength instead of time, and the optimization variables are selected to achieve piecewise constant curvature. The basic concept is to develop a road kinematic model with curvature as a state variable, and then regulate the derivative of the curvature using \( \ell_1 \)-norm regularization. This approach is capable of rendering a sparse vector for the derivative of curvature.

This article is organized as follow. Section IV presents the vehicle kinematics. Based on the vehicle kinematics and a few assumptions, the road kinematic model is presented in the Section V. With the road kinematic model, the problem is formulated in the optimization framework in the Section VI. The discussion of roadway data acquisition is presented in the Section VII. The experimental results and data storage are discussed in the Sections VIII and IX. The article concludes in the Section X.

IV. VEHICLE KINEMATICS

The kinematic model of a vehicle traveling without slip on a planar surface is

\[
\begin{align*}
\dot{p}_x(t) &= V(t) \cos(\psi(t)) \\
\dot{p}_y(t) &= V(t) \sin(\psi(t)) \\
\dot{\psi}(t) &= \omega(t)
\end{align*}
\]

where \([p_x, p_y]\) denotes position, \( \psi \) is the yaw angle of the vehicle relative to the \( x \)-axis, and the inputs to the kinematic model are the speed \( V \) and the angular rate \( \omega \).

V. ROAD MODEL

Given that the roadway is locally planar and intended to be traveled by vehicles, the kinematic model of eqns. (1-3) is the foundation of the roadway model. Since the road shape is not evolving versus time but as a function of arclength \( s \), some modifications are necessary to derive the road kinematic model.

Noting that \( \dot{p}_x(t) = dp_x(t)/dt \), \( \dot{p}_y(t) = dp_y(t)/dt \), \( \dot{\psi}(t) = d\psi(t)/dt \), and that \( ds = V(t)dt \), eqns. (1-3) can be rewritten as

\[
\begin{align*}
dp_x(t) &= \cos(\psi(t)) \ ds \\
dp_y(t) &= \sin(\psi(t)) \ ds \\
d\psi(t) &= \kappa(t) \ ds \\
d\kappa(t) &= \eta(t) \ ds
\end{align*}
\]

where \( \kappa = \omega(t)/V(t) \) represents curvature and \( \eta \) is the derivative of the curvature \( \kappa \) with respect to arclength \( s \).
The equivalent discrete-time model by Euler integration is

\[ p_x(k+1) = p_x(k) + \Delta s(k) \cos \psi(k) \]  \hspace{1cm} (8)
\[ p_y(k+1) = p_y(k) + \Delta s(k) \sin \psi(k) \]  \hspace{1cm} (9)
\[ \psi(k+1) = \psi(k) + \kappa(k) \Delta s(k) \]  \hspace{1cm} (10)
\[ s(k+1) = s(k) + \kappa(k) \Delta s(k) \]  \hspace{1cm} (11)
\[ \kappa(k+1) = \kappa(k) + \beta(k) \]  \hspace{1cm} (12)

where \( s(k) \), \( \kappa(k) \) are the arclength and the curvature of the
k-th point on the curve, \( \beta(k) \) is the increment of \( \kappa(k) \) at
point \( k \) and \( \beta(k) = \eta \Delta s(k) \), \( \Delta s(k) = V(k)T_s \) where \( T_s \) is
the integration time step.

Eqs. (8-12) define the road kinematic model with the
state vector:

\[ x(k) = [p_x, p_y, \psi, s, \kappa]_k. \] \hspace{1cm} (13)

The model inputs are \( \Delta s(k) \) and \( \beta(k) \). An important point to
notice here is that the curvature \( \kappa \) is exclusively controlled by \( \beta \) which allows the \( \ell_1 \)-norm regularization approach to yield
trajectories with piecewise constant curvature. Note that this
road kinematic model is accurate only when the change of
\( \psi(k) \), equivalently \( \kappa(k) \Delta s(k) \), is sufficiently small. Physically,
the separation between data points should be small relative to the roadway radius of curvature \( (\Delta s(k) \ll \frac{1}{\kappa(k)}). \)

VI. PROBLEM FORMULATION

The road curve fitting problem can be formulated as a non-
linear constrained optimization with an \( \ell_1 \)-norm regularization
in the following way:

\[ \mathbf{P}_0 : \min_{x(0), \beta(k), \Delta s(k)} \sum \| y(k) - W(k) \|^2_2 + \lambda \| b \|^1 \] \hspace{1cm} (14)

subject to the road dynamic model constraint defined in eqns.
(8-12) and

\[ -\Delta s(k) \leq 0, k = 0, \ldots, N - 1 \] \hspace{1cm} (15)

where \( y(k) = [p_x(k), p_y(k)] \) is the curve fit trajectory location at the k-th arc length to eqns. (8-12)
\( b = [\beta(0), \ldots, \beta(N - 1)] \), \( W(k) \in D = \{N_k, E_k\}_{k=0} \)
is the measured position along the road, and \( \lambda \) is a parameter
that weights the tradeoff between trajectory smoothness and
accuracy of the data fit. The inequality constraint (15)
ensures that the increments of arc length are non-negative.

It is well known that \( \ell_1 \)-norm regularization is able to render sparse results [10], [11], in the sense that most of
the elements of the vector \( b \) are zero. Because the non-zero
elements in the sparse vector \( b \) represent the discontinuities in curvature \( \kappa(k) \), the constraint of piecewise constant curvature
can be enforced in the above formulation. Optimization
problem \( \mathbf{P}_0 \) finds the best tradeoff between the road curve
data fit and the number of changes in the curvature.

In order to solve the optimization problem \( \mathbf{P}_0 \), there is
a standard trick to convert the \( \ell_1 \)-norm into a vector dot product
by introducing additional inequality constraints. Thus, by
defining the positive vector \( z = [z(0), \ldots, z(N-1)]^T \), the
optimization problem (14) can be rephrased as

\[ \min_{x(0), \beta(k), \Delta s(k)} \sum \| y(k) - W(k) \|^2_2 + \lambda^T z \] \hspace{1cm} (16)

subject to the constraints in eqns. (8-12), (15) and

\[ -z(k) \leq \beta(k) \leq z(k), k = 0, \ldots, N - 1 \] \hspace{1cm} (17)

If a stacked vector is defined as:

\[ \mathbf{X} = [x(0), \ldots, x(N), b, z, \Delta S] \]

where \( \Delta S = [\Delta s(0), \ldots, \Delta s(N - 1)] \), then the optimization
problem can be further written into a more compact form:

\[ \mathbf{P}_1 : \min_{\mathbf{X}} f(\mathbf{X}) \]

\[ s.t. \ q(\mathbf{X}) = 0 \]
\[ c(\mathbf{X}) \leq 0 \]

where \( f(\mathbf{X}) \) is the cost function in (16), \( q(\mathbf{X}) \) is the equality
constraint in eqns. (8-12), and \( c(\mathbf{X}) \) is the inequality
constraints in (15) and (17).

Now the optimization problem is in a standard form and
thus can be solved by classical interior-point algorithms
(such as “fmincon” in MATLAB®). To get a better result,
we can solve the optimization problem iteratively [10]:

- Solve the optimization problem \( \mathbf{P}_1 \) using \( \lambda = \lambda_1 \) to get
  a sparse vector \( b = [\beta(0), \ldots, \beta(N - 1)] \)
- Add additional constraints: \( \beta(k) = 0 \) for \( k \in \mathcal{K} = \{ k | \beta(k) \leq \beta_0 \} \) and choose a different \( \lambda = \lambda_2 \).
- Then solve the new optimization problem again.

Typically, the problem converges in two iterations. In this
paper, \( \beta_0 \) is chosen to be \( 10^{-4} \). Other more refined methods,
such as weighted \( \ell_1 \) minimization [12], can also be applied
here to solve \( \mathbf{P}_0 \).

VII. ROADWAY DATA

The following section demonstrates the algorithm. To
attain the lane-level accuracy, a set of accurate roadway data
is required. At the time the paper was written, a smoothing
algorithm had not yet been implemented; therefore, the data
that we use is the result of real-time vehicle state estimation.
A test vehicle was instrumented with a Carrier Phase Differen-
tial GPS aided Inertial Navigation System (INS) [13]–[17]
that is capable of achieving vehicle positioning accuracies
better than 0.1m while the vehicle is maneuvering at highway
speeds. The INS state vector is saved at a 30 Hz rate. Only
the INS position estimates are used in the approach that
follows. To acquire data for the examples portion of this
paper, we drove the test vehicle along a roadway that was
of interest.

The vehicle drove at a maximum speed of 60 miles/h and
the sampling rate is 30Hz. Therefore, \( \Delta s(k) \) is less that
0.81 meter which is small relative to the (minimum) radius of
curve \( (1/\kappa(k)) \) experienced in the data set. Therefore, the
road kinematic model of eqns. (8-12) is applicable. The data
is obtained for a trajectory merging onto Highway 215 at Fair
Isle Dr. in Riverside CA. This dataset contains 800 points
(376 m in arc length).

This data acquisition methodology results in two sources of
roadway position estimation error. The first source is the
error in the INS estimate of the vehicle position. The second
source is the inability of the driver to exactly follow the
roadway centerline. While the INS vehicle position error is small (<0.1m) and uncorrelated for lags greater than a few seconds. The driver error is bounded only by one half the lane width, but may be biased and correlated over short periods of time. Nonetheless, the curve fit results show excellent performance and will only be further enhanced when the input data \( D \) is the result of a smoothed data fusion process.

Eventually it is expected [18], [19] that large volumes of vehicle position data will become available when volunteer drivers allow their GPS enabled cell phones to be used as probes. While this data source will introduce new challenges (e.g., detection of lane changes, detection of erroneous data, estimation of phone position relative to the vehicle center), it also highlights the fact that in the near future large volumes of lane data will become available and reliable methods for automatically extracting high-accuracy maps will be required. Existing and emerging roadway geometry data collection methods are discussed in [5].

VIII. EXPERIMENTAL RESULTS

Trial 1 used \( \lambda_1 = 150 \) in the first iteration of the optimization and \( \lambda_2 = 180 \) in the second iteration of the optimization. The results are shown in Fig. 1. The upper left subplot shows the trajectory drawn on the tangent plane. The coordinates of the road have been shifted to be centered at zero. The red asterisks indicate the points where discrete changes in curvature occur (i.e., nodes, see Section IX). The travel direction is from left to right. The top right subplot shows the trajectory yaw angle which is continuous and piecewise linear. The blue curve in the right subplot in the second row shows the curvature \( \kappa \), which clearly exhibits the piecewise constant property. For the comparison purpose, we also plot the curvature computed from the spline curve fit method [8] using cubic Hermite splines. For spline method 1 (green), we choose the spline knots evenly every 50 meters. For spline method 2 (red), we choose the same knots at the points of curvature discontinuity identified by the \( \ell_1 \) optimization. It is clear that in both cases, the spline curvature is not piecewise constant. In addition the variation of the spline curvature is significantly larger than is necessary to fit the \( D \). Finally, the spline curvature has the incorrect sign at some locations.

The bottom right subplot shows the change in curvature, i.e. \( \beta \), which is an impulse train as desired. The left subplot in the second row shows the curve fit error that is computed pointwise as

\[
e(k) = \hat{N}(k) \cdot (W(k) - y(k))
\]

where \( \hat{N}(k) = [\sin \psi(k), -\cos \psi(k)]^T \) is the (right) normal vector to the fitted curve at the \( k \)-th point. Therefore, \( e(k) \) is positive if \( W(k) \) is on the right hand side of the fitted curve and negative if \( W(k) \) is on the left hand side of the fitted curve. It is clear that the magnitude of the curve fit error is less than 5 cm at most locations and never larger than 10 cm. In addition, it is clear that the road has two portions. The first portion of the road has high curvature while the second portion has lower curvature. The dividing line is marked in the subplot by the red dash vertical line that corresponds to the fourth curvature change. The mean square error (MSE) for each portion of the road is shown in row one of Table I. Table I also shows summary results of a second trial using \( \lambda_1 = 20 \) and \( \lambda_2 = 80 \). For both trials, the MSE is greater in the first portion of the road than the second portion. This is reasonable given that the driver has greater difficulty keeping the vehicle on the centerline of the road when the road has significant curvature. Table I also shows that the different choice of weights \( \lambda \) did not drastically change the accuracy nor the number of required nodes. The bottom left subplot displays the unbiased error auto-correlation function showing that the spatial correlation of the error drops below \( 4 \times 10^{-3} \) in a few meters. This illustrates that the fit error, as mentioned in Section VII, is only correlated over a short time periods.

The algorithm optimally selects the number and the locations of curvature change points jointly to minimize the cost function defined in (16). Therefore, this algorithm eliminates the need of any pre-processing to search for curvature change. Moreover, there are only two parameters to tune (\( \lambda \) and \( \beta_0 \)). Regarding computational complexity, the optimization algorithm is very efficient (done in less than 2 minutes) for dataset presented herein. However, due to the nonlinearity in eqns. (8)-(12), the problem is not a convex optimization problem. Hence, sometimes the optimization algorithm converges slowly. But by choosing a different \( \lambda \) the problem can usually be solved quickly.

IX. DATA STORAGE AND REALTIME USAGE

At the completion of the optimization, we obtain the optimal curve as a sequence of points \( T = \{x(k), \beta(k)\}_{k=0}^N \) (see eqn. (13)). Only the nodes \( N = \{x(k + 1) | \beta(k) \neq 0\} \) need to be stored as they are sufficient to exactly represent \( T \). To clarify the presentation, we use \( n \) to count over the nodes in \( N \). For node \( n \), we denote its position as \( P_n = [P_{x}(n), P_{y}(n)]^T \).

Applications require a parameterized function \( \gamma(s) \) that computes the trajectory \( T \) as a function of arclength \( s \in [0, s_N] \). Using only the nodes \( N \), the procedure is as follows:

1) Given \( s \in [0, s_N] \), find \( n \) such that \( s \in [s_n, s_{n+1}] \).

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>OPTIMIZATION PERFORMANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>MSE (( \times 10^{-2} ))</td>
</tr>
<tr>
<td>s \in [0,115]m</td>
<td>s \in [115,375]m</td>
</tr>
<tr>
<td>1:</td>
<td>1.42</td>
</tr>
<tr>
<td>2:</td>
<td>0.49</td>
</tr>
</tbody>
</table>

\(^1\)The discrete changes in the curve fit error are caused by the realtime Kalman filter corrections at GPS measurement times GPS. These would be absent if smoothed data were available.
2) If $\kappa_n = 0$, then this segment is a line; therefore,
\[
\gamma(s) = P_n + \mu T_n
\]
where $\mu = \frac{s - s_n}{s_{n+1} - s_n}$ and $T_n = P_{n+1} - P_n$. The curvature is zero and the direction is $\psi_n$ which should match the four quadrant arctangent of $T$.

3) If $\kappa \neq 0$, then the segment is a circle. The center of the circle can be computed by either
\[
P_c = P_n + R_n \hat{N}_n
\]
\[
P_c = P_{n+1} + R_n \hat{N}_{n+1}
\]
where $R_n = \frac{1}{|\kappa_n|}$ is the radius of curvature and $\hat{N}_n = -\text{sign}(\kappa_n)[\sin \psi_n, - \cos \psi_n]^T$ is the inward pointing unit normal. Both equations should yield exactly the same center location, which is useful as a check. The direction of the tangent to the circle at $s$ is
\[
\psi(s) = \psi_n + \kappa_n(s - s_n).
\]
The location on the circle at arclength $s$ is
\[
\gamma(s) = P_c - R_n \hat{N}(s)
\]
where $\hat{N}(s) = -\text{sign}(\kappa_n)[\sin \psi(s), - \cos \psi(s)]^T$.

The above computation procedure is straightforward and exactly represents $T$. The resulting $\gamma(s)$ is available for real-time applications.

In comparison with traditional roadway representations that store only $[s_n, P_n]$ and use linear connections, the approach above requires storage of $[s_n, P_n, \psi_n, \kappa_n]$ which is two extra parameters. The tradeoff is that a more accurate roadway representation is obtained. In addition, it may turn out that with the improved representation, the nodes can be more widely spaced, allowing for the overall memory requirements per lane to stay constant.

X. CONCLUSIONS AND FUTURE RESEARCH

This article has discussed the problem of generating roadway trajectory curve fits that accurately match available data while enforcing roadway characteristics known from the design principles (i.e., piecewise constant curvature). The article includes experimental results and discussion of the straightforward analytic trajectory representation in terms of nodes.

This work could be extended in several different directions. First, ultimately, the curve fit procedure must work in three dimensions. A challenge in that extension is that the vertical and lateral curvature changes are independent. Second, the INS or smoothing process maintains an estimate of the position error covariance matrix at each time instant. This error covariance matrix could be incorporated as a time varying weight into the trajectory optimization procedure. This article has only used the measured position data in the cost function. The INS or smoothed trajectory also has other information (e.g., velocity, attitude, angular rate) that might be useful.

XI. ACKNOWLEDGMENTS

This material is based upon work supported by the California Department of Transportation under Contract No. 65A0311 and the DOT Federal Highway Administration (FHA) Agency Award No: DTFH61-09-C-00018. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsors.

REFERENCES